

11/3/

#2



OIPE

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/10/032,658

TIME: 09:51:59

Input Set : N:\Crf3\RULE60\10032658.raw

Output Set: N:\CRF3\02062002\J032658.raw

ENTERED

SEQUENCE LISTING

3 (1) GENERAL INFORMATION:

5 (i) APPLICANT: Graham, Laurie A.

6 Liou, Yih-Cherng

7 Walker, Virginia K.

8 Davies, Peter L.

10 (ii) TITLE OF INVENTION: Tenebrio Antifreeze Proteins

12 (iii) NUMBER OF SEQUENCES: 22

14 (iv) CORRESPONDENCE ADDRESS:

15 (A) ADDRESSEE: Townsend and Townsend and Crew LLP

16 (B) STREET: Two Embarcadero Center, Eighth Floor

17 (C) CITY: San Francisco

18 (D) STATE: California

19 (E) COUNTRY: USA

20 (F) ZIP: 94111-3834

22 (v) COMPUTER READABLE FORM:

23 (A) MEDIUM TYPE: Floppy disk

24 (B) COMPUTER: IBM PC compatible

25 (C) OPERATING SYSTEM: PC-DOS/MS-DOS

26 (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

28 (vi) CURRENT APPLICATION DATA:

C--> 29 (A) APPLICATION NUMBER: US/10/032,658

C--> 30 (B) FILING DATE: 02-Jan-2002

31 (C) CLASSIFICATION:

34 (vii) PRIOR APPLICATION DATA:

35 (A) APPLICATION NUMBER: US/08/882,907

36 (B) FILING DATE: 26-JUN-1997

38 (viii) ATTORNEY/AGENT INFORMATION:

39 (A) NAME: Weber, Kenneth A.

40 (B) REGISTRATION NUMBER: 31,677

41 (C) REFERENCE/DOCKET NUMBER: 016252-002100US

43 (ix) TELECOMMUNICATION INFORMATION:

44 (A) TELEPHONE: (415) 576-0200

45 (B) TELEFAX: (415) 576-0300

48 (2) INFORMATION FOR SEQ ID NO: 1:

50 (i) SEQUENCE CHARACTERISTICS:

51 (A) LENGTH: 12 amino acids

52 (B) TYPE: amino acid

53 (C) STRANDEDNESS:

54 (D) TOPOLOGY: linear

56 (ii) MOLECULE TYPE: peptide

59 (ix) FEATURE:

60 (A) NAME/KEY: Peptide

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/032,658

DATE: 02/06/2002

TIME: 09:51:59

Input Set : N:\Crf3\RULE60\10032658.raw

Output Set: N:\CRF3\02062002\J032658.raw

```

61      (B) LOCATION: 1..12
62      (D) OTHER INFORMATION: /note= "consensus 12 amino acid
63 repeating motif"
66      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 1:
W--> 68      Cys Thr Xaa Ser Xaa Xaa Cys Xaa Xaa Ala Xaa Thr
69      1          5          10
72 (2) INFORMATION FOR SEQ ID NO: 2:
74      (i) SEQUENCE CHARACTERISTICS:
75          (A) LENGTH: 68 base pairs
76          (B) TYPE: nucleic acid
77          (C) STRANDEDNESS: single
78          (D) TOPOLOGY: linear
W--> 80      (ii) MOLECULE TYPE: DNA
83      (ix) FEATURE:
84          (A) NAME/KEY: -
85          (B) LOCATION: 1..68
86          (D) OTHER INFORMATION: /note= "conserved N-terminal region of
87 thermal hysteresis protein (THP)
88 isoforms"
91      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 2:
93 TGCAC TGGGG STGCTGATTG YACTAGTTGT ACAGVWGCAT GCACTGGTTG TGGAARYTGT      60
95 CCAAATGC                                                                68
98 (2) INFORMATION FOR SEQ ID NO: 3:
100     (i) SEQUENCE CHARACTERISTICS:
101         (A) LENGTH: 16 amino acids
102         (B) TYPE: amino acid
103         (C) STRANDEDNESS:
104         (D) TOPOLOGY: linear
106     (ii) MOLECULE TYPE: peptide
109     (ix) FEATURE:
110         (A) NAME/KEY: Peptide
111         (B) LOCATION: 1..16
112         (D) OTHER INFORMATION: /note= "consensus 16 amino acid
113 N-terminal motif for YL-1, YL-2, YL-3
114 and YL-4"
117     (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 3:
W--> 119     Xaa Cys Thr Xaa Xaa Xaa Xaa Cys Thr Xaa Cys Thr Xaa Xaa Cys Thr
120     1          5          10          15
123 (2) INFORMATION FOR SEQ ID NO: 4:
125     (i) SEQUENCE CHARACTERISTICS:
126         (A) LENGTH: 24 amino acids
127         (B) TYPE: amino acid
128         (C) STRANDEDNESS:
129         (D) TOPOLOGY: linear
131     (ii) MOLECULE TYPE: peptide
134     (ix) FEATURE:
135         (A) NAME/KEY: Peptide
136         (B) LOCATION: 1..24
137         (D) OTHER INFORMATION: /note= "N-terminal amino acid sequence

```

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/032,658

DATE: 02/06/2002

TIME: 09:51:59

Input Set : N:\Crif3\RULE60\10032658.raw

Output Set: N:\CRF3\02062002\J032658.raw

138 of YL-1, YL-2, YL-3 and YL-4"

140 (ix) FEATURE:

141 (A) NAME/KEY: Modified-site

142 (B) LOCATION: 1

143 (D) OTHER INFORMATION: /product= "OTHER"

144 /note= "Xaa = Gln or His"

146 (ix) FEATURE:

147 (A) NAME/KEY: Modified-site

148 (B) LOCATION: 5

149 (D) OTHER INFORMATION: /product= "OTHER"

150 /note= "Xaa = Ala or Gly"

152 (ix) FEATURE:

153 (A) NAME/KEY: Modified-site

154 (B) LOCATION: 13

155 (D) OTHER INFORMATION: /product= "OTHER"

156 /note= "Xaa = Ala, Asp or Gly"

158 (ix) FEATURE:

159 (A) NAME/KEY: Modified-site

160 (B) LOCATION: 20

161 (D) OTHER INFORMATION: /product= "OTHER"

162 /note= "Xaa = Asn or Ser"

165 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 4:

W--> 167	Xaa Cys Thr Gly Xaa Ala Asp Cys Thr Ser Cys Thr Xaa Ala Cys Thr
168	1 5 10 15

W--> 170	Gly Cys Gly Xaa Cys Pro Asn Ala
171	20

174 (2) INFORMATION FOR SEQ ID NO: 5:

176 (i) SEQUENCE CHARACTERISTICS:

177 (A) LENGTH: 61 base pairs

178 (B) TYPE: nucleic acid

179 (C) STRANDEDNESS: single

180 (D) TOPOLOGY: linear

W--> 182 (ii) MOLECULE TYPE: DNA

185 (ix) FEATURE:

186 (A) NAME/KEY: -

187 (B) LOCATION: 1..61

188 (D) OTHER INFORMATION: /note= "conserved C-terminal region of
189 thermal hysteresis protein (THP)

190 isoforms"

193 (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 5:

195 CTCAACCAAC TGTTACAAAG CTACAGCCTG TACCAATTCA WCAGGATGTC CCGGACATTA 60

197 R 61

200 (2) INFORMATION FOR SEQ ID NO: 6:

202 (i) SEQUENCE CHARACTERISTICS:

203 (A) LENGTH: 32 base pairs

204 (B) TYPE: nucleic acid

205 (C) STRANDEDNESS: single

206 (D) TOPOLOGY: linear

W--> 208 (ii) MOLECULE TYPE: DNA

RAW SEQUENCE LISTING

PATENT APPLICATION: US/10/032,658

DATE: 02/06/2002

TIME: 09:51:59

Input Set : N:\Cr3\RULE60\10032658.raw

Output Set: N:\CRF3\02062002\J032658.raw

```

211      (ix) FEATURE:
212          (A) NAME/KEY: -
213          (B) LOCATION: 1..32
214          (D) OTHER INFORMATION: /note= "PCR primer"
217      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 6:
219 CATATGCATA TGCAATGCAC TGGGGGTGCT GA
222 (2) INFORMATION FOR SEQ ID NO: 7:
224      (i) SEQUENCE CHARACTERISTICS:
225          (A) LENGTH: 32 base pairs
226          (B) TYPE: nucleic acid
227          (C) STRANDEDNESS: single
228          (D) TOPOLOGY: linear
W--> 230      (ii) MOLECULE TYPE: DNA
233      (ix) FEATURE:
234          (A) NAME/KEY: -
235          (B) LOCATION: 1..32
236          (D) OTHER INFORMATION: /note= "PCR primer"
239      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 7:
241 AAGCCTAAGC TTTTAATGTC CGGGACATCC TG
244 (2) INFORMATION FOR SEQ ID NO: 8:
246      (i) SEQUENCE CHARACTERISTICS:
247          (A) LENGTH: 18 base pairs
248          (B) TYPE: nucleic acid
249          (C) STRANDEDNESS: single
250          (D) TOPOLOGY: linear
W--> 252      (ii) MOLECULE TYPE: DNA
255      (ix) FEATURE:
256          (A) NAME/KEY: -
257          (B) LOCATION: 1..18
258          (D) OTHER INFORMATION: /note= "internal downstream facing
259 sequencing primer"
262      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 8:
264 AAAGACTGTT TTGAAGCC
267 (2) INFORMATION FOR SEQ ID NO: 9:
269      (i) SEQUENCE CHARACTERISTICS:
270          (A) LENGTH: 18 base pairs
271          (B) TYPE: nucleic acid
272          (C) STRANDEDNESS: single
273          (D) TOPOLOGY: linear
W--> 275      (ii) MOLECULE TYPE: DNA
278      (ix) FEATURE:
279          (A) NAME/KEY: -
280          (B) LOCATION: 1..18
281          (D) OTHER INFORMATION: /note= "internal upstream facing
282 sequencing primer"
285      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 9:
287 TTCAAAACAG TCTTTTGA
290 (2) INFORMATION FOR SEQ ID NO: 10:
292      (i) SEQUENCE CHARACTERISTICS:

```

RAW SEQUENCE LISTING

DATE: 02/06/2002

PATENT APPLICATION: US/10/032,658

TIME: 09:51:59

Input Set : N:\Crf3\RULE60\10032658.raw

Output Set: N:\CRF3\02062002\J032658.raw

```

293      (A) LENGTH: 559 base pairs
294      (B) TYPE: nucleic acid
295      (C) STRANDEDNESS: single
296      (D) TOPOLOGY: linear
298      (ii) MOLECULE TYPE: cDNA
301      (ix) FEATURE:
302          (A) NAME/KEY: -
303          (B) LOCATION: 1..559
304          (D) OTHER INFORMATION: /note= "YL-1 thermal hysteresis protein
305 (THP) cDNA from Tenebrio molitor"
307      (ix) FEATURE:
308          (A) NAME/KEY: CDS
309          (B) LOCATION: 37..375
310          (D) OTHER INFORMATION: /product= "YL-1"
313      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 10:
315 TAAACAGCGA GATAACAAC AATACTACAT AAAACT ATG GCG TTC AAA ACG TGT      54
316                                     Met Ala Phe Lys Thr Cys
317                                     1           5
319 GGT TTT TCA AAA AAA TGG TTA GTA ATA GCA GTT ATA GTT ATG TGT TTG      102
320 Gly Phe Ser Lys Lys Trp Leu Val Ile Ala Val Ile Val Met Cys Leu
321          10           15           20
323 TGT ACC GAG TGT TAT TGC CAC TGC ACT GGG GGT GCT GAT TGT ACT AGT      150
324 Cys Thr Glu Cys Tyr Cys His Cys Thr Gly Gly Ala Asp Cys Thr Ser
325          25           30           35
327 TGT ACA GAT GCA TGC ACT GGT TGT GGA AAT TGT CCA AAT GCA CAT ACG      198
328 Cys Thr Asp Ala Cys Thr Gly Cys Gly Asn Cys Pro Asn Ala His Thr
329          40           45           50
331 TGT ACC GAT TCC AAA AAT TGT GTC AAG GCA GCA ACA TGT ACT GGA TCT      246
332 Cys Thr Asp Ser Lys Asn Cys Val Lys Ala Ala Thr Cys Thr Gly Ser
333          55           60           65           70
335 ACA AAA TGT AAT ACC GCC AGG ACG TGT ACA AAC TCA AAA GAC TGT TTT      294
336 Thr Lys Cys Asn Thr Ala Arg Thr Cys Thr Asn Ser Lys Asp Cys Phe
337          75           80           85
339 GAA GCC AAA ACA TGT ACT GAC TCA ACC AAC TGT TAC AAA GCT ACA GCC      342
340 Glu Ala Lys Thr Cys Thr Asp Ser Thr Asn Cys Tyr Lys Ala Thr Ala
341          90           95           100
343 TGT ACC AAT TCA ACA GGA TGT CCC GGA CAT TAAGTTTTTC TATTGTCAAC      392
344 Cys Thr Asn Ser Thr Gly Cys Pro Gly His
345          105           110
347 AATAATAAAA CACACTTACT GTTATCTTAG CTAAACATA ATTGTAAGCT CACTCTGTTT      452
349 TGTATCCTAT CTGTCTCTGC CTCCGAAGGA TGATAATTTT GTACTGGGAG CGAAAGGTTT      512
351 ATCCGACAAT AATAAACTAA AATAATTGAT ATAAAAAAAA AAAAAAA      559
354 (2) INFORMATION FOR SEQ ID NO: 11:
356      (i) SEQUENCE CHARACTERISTICS:
357          (A) LENGTH: 112 amino acids
358          (B) TYPE: amino acid
359          (D) TOPOLOGY: linear
361      (ii) MOLECULE TYPE: protein
363      (xi) SEQUENCE DESCRIPTION: SEQ ID NO: 11:

```

VERIFICATION SUMMARY

PATENT APPLICATION: US/10/032,658

DATE: 02/06/2002

TIME: 09:52:00

Input Set : N:\Crf3\RULE60\10032658.raw

Output Set: N:\CRF3\02062002\J032658.raw

L:29 M:220 C: Keyword misspelled or invalid format, [(A) APPLICATION NUMBER:]
L:30 M:220 C: Keyword misspelled or invalid format, [(B) FILING DATE:]
L:68 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:1
L:80 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=2
L:119 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:3
L:167 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:170 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:4
L:182 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=5
L:208 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=6
L:230 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=7
L:252 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=8
L:275 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=9
L:853 M:246 W: Invalid value of Alpha Sequence Header Field, [MOLECULE TYPE:], SeqNo=21
L:892 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22
L:895 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:22